DOCUMENT RESUME

ED 117 168 TM 005 023

AUTHOR Foreman, Dale I.
TITLE SCORE: A Flexible System for School Curriculum Objective-Referenced Evaluation.
PUB DATE Apr 75
NOTE 10p.; Paper presented at the Annual Meeting of the American Educational Research Association (Washington, D.C., March 30-April 3, 1975)

EDRS PRICE MF-$0.76 HC-$1.58 Plus Postage
DESCRIPTORS Computer Programs; *Criterion Referenced Tests; *Educational Objectives; Elementary Secondary Education; *Item Banks; Language Arts; Mathematics; Reading; Sciences; Social Studies; Student Evaluation; Test Construction
IDENTIFIERS *School Curriculum Objective Referenced Evaluation; SCORE

ABSTRACT

The School Curriculum Objective-Referenced Evaluation (SCORE) system is an item bank that can be effectively and inexpensively used to develop criterion referenced tests which reflect an institution's educational objectives. SCORE also allows for test construction, formatting, and reporting. Flexible in design, SCORE educational objectives are organized in a three-level hierarchy of goals, intermediate objectives, and performance objectives. SCORE attempts to test each educational objective in the most appropriate way and may require special equipment. Student performance reports are shared with the student, parent, school, and school district. (BJG)

****
* Documents acquired by ERIC include many informal unpublished *
* materials not available from other sources. ERIC makes every effort *
* to obtain the best copy available. Nevertheless, items of marginal *
* reproducibility are often encountered and this affects the quality *
* of the microfiche and hardcopy reproductions ERIC makes available *
* via the ERIC Document Reproduction Service (EDRS). EDRS is not *
* responsible for the quality of the original document. Reproductions *
* supplied by EDRS are the best that can be made from the original. *
****

ED117168

SCORE: A Flexible System for School Curriculum Objective-Referenced Evaluation

Dale I. Foreman
Westinghouse Learning Corporation

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
NATIONAL INSTITUTE OF EDUCATION
THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

There is a trend today toward refocusing the emphasis in testing and evaluation on objective-referenced (or criterion-referenced) methods of test construction and measurement. This trend can be seen in the new testing programs of most publishers to the extent that even the norm-referenced Stanford 73 test now has a published list of objectives that make up the test.

Other evidence of this thrust, and no doubt a major causal factor, is the emphasis of state legislatures to make the educational systems (schools) accountable. This push requires the movement toward defining goals and setting objectives for education, and the evaluation of programs based upon these goals and objectives. Many teacher groups have been involved in workshops geared toward writing behavioral objectives, in order to prepare for the legislative-inspired accountability.

Testing companies have responded to these demands in many different ways. These include (1) banking educational objectives either in text or on the computer; (2) developing "criterion-referenced" tests with items matched to objectives; (3) banking test items which are referenced to objectives; and (4) developing flexible objective-referenced testing systems which are computer based.

School Curriculum Objective-Referenced Evaluation (SCORE) is an example of a flexible system of objective-referenced (criterion-referenced) evaluation. This paper will discuss in detail the SCORE system, pointing out its unique features, as well as advantages over testing programs that are more limited.

TM005 023

The basic philosophy behind the SCORE system is, that although many of the basic skills that every student is required to learn are the same or very similar, there are differences in both the scope of the curriculum and the sequence in which it is presented. Priorities differ from school district to district, and emphasis and sequencing of selected vital objectives also varies.

SCORE is designed to make testing directly applicable to and reflective of each school district's established instructional program. This responsiveness allows testing to take place on any set of objectives at any point in time.

SCORE was designed to be a dynamic system of evaluation, including objectives and item banks, test creation and formatting, plus a reporting system directed at the needs of students and teachers. The design of the system allows the SCORE curriculum staff to add to, modify, or delete portions of the data bank in response to curricular trends without destroying the basic structure.

Because of this flexibility in design, the data bank can be frequently reviewed and revised. New objectives catalogs are created every two to three years. These catalog revisions normally include (1) addition of objectives which cover new curricular material or material not previously considered important enough to include; (2) deletion of objectives which have become obsolete or are not in demand by the users of the system; and (3) reorganization whenever there is evidence that it will increase the usefulness of the catalog.

Test items can also be reviewed periodically and revised or deleted if necessary. Many times, the revision is based upon input from the users.

SCORE is unique. First it is objectives. SCORE is a bank of about 26,000 educational objectives. The objectives have been stored in a computer data bank using an easy access coding scheme which allows them to be retrieved via a simple index number. The objectives cover grades K-9 in reading/language arts, mathematics, science, and social studies. They are provided in printed catalog form for users. The catalogs are printed in three volumes; primary (K-3), intermediate (4-6), and junior high (7-9).

The objectives in the catalogs have been carefully written and organized into a three-level hierarchy. The three levels of this hierarchy are goals, intermediate objectives, and performance objectives. These levels provide an organizational tool. They also allow the user easy access to objectives of interest.

The goal is a very broad statement of an area within the curriculum. For example:

"The student will demonstrate an understanding of man's adaptation."

In general, the goals are not behaviorally stated. They represent concepts that cut across many grades in differing detail and breadth from grade to grade. They often relate well to district program goals.

The intermediate objective is a more narrow subset of the goal. For example:

"The student will demonstrate an understanding of the adaptation of early man to his environment."

Sometimes the intermediate objective is stated in behavioral terms, but it is not necessary. The objectives can be thought of as course or unit divisions which organize or elaborate the concept status in the goal.

The final level of objectives is the performance objective. It is a very specific statement of expected performance which the student can exhibit. For example:

"The student will be able to identify or select one or more examples of natural selection occurring in man."

The performance objective is always stated in behavioral terms. It describes a student behavior that is observable. These behaviors may include written response, oral response, or overt action. These objectives are the basic prescriptions that are used to define the test items.

The SCORE system includes objectives which will accommodate items for all levels of Bloom's taxonomy. This gives the user the capability of tapping all levels of cognative ability.

SCORE is items. Currently there are about 30,000 test items in the data banks. Each performance objective has attached to it a pool of items. The number of items per objective varies greatly, ranging from one to 15 or 20. Some objectives which are definitional in nature have only one or two items, while other objectives have 20 to 25 items in the bank.

The items in the bank are written in several different formats. Most of them are multiple-choice, but there are also items that require single-word responses, written-passage responses, and oral responses (requiring individual administration). Some test items call for the use of special equipment, such as tape recordings, science apparatus, aerial photographs and so forth. This flexibility makes it possible to test each objective in the most appropriate way so that the results are easily interpretable.

Within each objective, grade, and item format cluster, the items are written so that they are, as nearly as possible, statistically equivalent. Every multiple choice item that is written for a given objective at the fifth grade level should be of approximately the same difficulty. In this way, one can get the same measure of performance on an objective no matter which item is selected from the data bank. If an item is found to be of greater or lesser difficulty than other items for the same objective, it is either revised, reassigned to another objective, or reassigned to a different grade.

The multiple-choice test items are carefully written so that each item can be diagnostic within itself. To the degree possible, each foil has a reason connected to its use. In this way, if a student selects an incorrect foil, the teacher can identify some possible reason for the mistake, thus making remediations much easier. This is particularly true in the area of mathematics and less true in the more amorphous areas such as social studies.

SCORE is a meaningful system of reporting. The reports are designed to give information to the student, the teacher, the parents, the school, and the district. All levels of reporting are done in terms of mastery of objectives. The mastery of an objective is determined by comparing the number of items answered correctly to a criterion that is pre-established by the school district. For example, suppose that a school district has a mastery criterion of 80 percent. That means that for a student to master an objective, he/she must answer 4 out of 5 or 5 out of 5 items correctly to demonstrate mastery of the objective.

Objective mastery is computed at both the intermediate and performance objective levels. Because there are several performance objectives under each intermediate objective, the confidence in reported mastery or non-mastery is greater at the intermediate level than at the performance objective level.

One thing must be kept in mind. Unless a representative sampling of performance objectives has been selected under a given intermediate objective, the user cannot claim mastery of the total universe defined by the intermediate objective. The selection of objectives for meaningful reporting is an art that should be approached with care and much thought.

The reports include:

(1) A class list report gives student data for each student in the class in terms of performance and intermediate objectives. This is reported in terms of mastery and non-mastery by objective, including a summary for each student. Also, for each objective there is a summary of the percent of students in the class who mastered it. This report is the teacher's chief tool.

(2) A parent/teacher conference guide is a verbal report which gives a list of intermediate objectives that the student has mastered and a list of those the student has not mastered. This is a report that is appropriate for the parent and student. It paints a clear picture of what is expected of the student and how much he/she still has to learn to complete the requirements.

(3) The building and district summaries are provided to give data on the levels of performance for each teacher/class in a building, as well as for each building in the district. These summaries are useful in defining major weaknesses in the curriculum. This could be a valuable guide in refocusing or refining curriculum and instruction. But not to provide evaluation of teacher performance.

(4) Finally, there is a special item analysis report which provides a record of each student's response on every item. If the question is answered incorrectly, the student's response is given. This helps the teacher diagnose the student's problem and allows him/her to group for instruction all students with the same problem.

SCORE is total involvement by teachers, administrators, and all school personnel. From the very beginning, the teachers are asked to form selection committees to peruse the objectives catalogs and prescribe the nature and format of their test. They select objectives, item formats, and the number of items per objective that they want on the test. They can also control the reading level of the items by adjusting the grade level requested.

The tests can be any length the school district wants. Several equivalent forms can be created so that pre- and post-testing can be done on the same objectives with many of the items being different.

After the objectives are selected, SCORE creates a test mock-up which the teacher committee reviews. The committee can suggest item revisions. Once the revisions have been made, a custom booklet is created for the district in the exact quantity requested. A SCORE staff of professionals

is involved in every step of the test's development. Every item receives at least two professional reviews. Close contact is maintained with school district personnel to guarantee acceptability of the test.

SCORE is a mechanized test creation system. The items that make up the district's customized test booklet are randomly selected from the SCORE data bank. A test selection booklet which has been filled out by the teacher committee is scanned to identify, on the computer, those objectives to be tested. The computer then accesses the item bank and randomly selects the number of items requested for each objective.

These items are then formatted by the computer onto pages so that each page is completely filled. Graphics are then added by the graphics arts division. Once the final draft test is approved, the test booklets are printed in either machine scorable form or in booklets which require answer sheets.

SCORE has been used in many different ways. They include (1) a pre-post testing for Title I evaluation of individual students; (2) a pre-test for student placement; (3) a post test for student evaluation and grading; (4) for program evaluation using matrix sampling procedures; and (5) for periodic progress monitoring, utilizing single objective mini-tests.

SCORE was not designed as a norm-referenced test. No attempt has been made to gather data on instruments for purposes of norming. This is due in large part to the fact that every instrument is different from every other instrument. SCORE does have, however, a mechanism for cumulating item data on each item in the bank. This includes item analysis information,

as well as point biserial correlations of each item with other items for the same objective if used together on a test. This data can provide information for the user or for the SCORE staff when considering revision or deletion.

In summary, SCORE is an extremely flexible and responsive system. It is designed to accommodate the needs of educators (particularly teachers) and students. The focus is on improving the collection and dissemination of meaningful information which will improve instruction and, hopefully, learning.